考虑可变环境因素的公共自行车短期需求预测模型 *

乔 健†，陈少博，何梦莹

(西北工业大学 管理学院，西安 710072)

摘 要：现有公共自行车短期需求预测模型忽视了不同环境因素影响用户需求的性质差异和可变环境因素的时间依赖性。将环境因素区分为已内化于需求的不变因素和需要单独考虑的可变因素，提出一种用图卷积神经网络(GCNN)捕获用户需求的非欧式空间依赖、用长短期记忆(LSTM)网络捕获用户需求和可变环境因素的时间依赖、通过向量拼接和全连接网络将可变环境因素的影响施加于用户需求的 GCNN-LSTM-E 模型。实验结果显示，GCNN-LSTM-E 模型在 1h 时间粒度下的预测性能最优，而且优于所有基准模型的预测性能。说明该模型的设计合理有效，1h 是最合适的时间粒度。

关键词：公共自行车；短期需求预测；图卷积神经网络；长短期记忆

中图分类号：TP183;U491.225　　doi: 10.19734/j.issn.1001-3695.2022.01.0024

Short-term demand prediction model for public bikes considering variable environmental factors

Qiao Jian†, Chen Shaobo, He Mengying

(*School of Management*, *Northwestern Polytechnical University*, *Xi'an Shaanxi* 710072, *China*)

Abstract: Existing short-term demand prediction models for public bikes ignored the difference in the nature of the impacts of different environmental factors on user demand as well as the temporal dependency of variable environmental factors. this paper distinguished environmental factors into the invariable factors that have been internalized into user demand and the variable factors that need to be considered separately, and then proposed a model called GCNN-LSTM-E in this paper. In the model, this paper used graph convolutional neural network (GCNN) to capture the non-euclidean spatial dependency of user demand, used long short-term memory (LSTM) network to capture the temporal dependencies of user demand and variable environmental factors, and applied vector concatenation and fully connected network to impose the influence of variable environmental factors on user demand. Experimental results showed that the GCNN-LSTM-E model has the best prediction performance and outperforms all benchmark models under 1h time granularity. It indicates that the design of the model is reasonable and effective, and 1h is the most appropriate time granularity.

Key words: public bike; short-term demand prediction; graph convolutional neural network; long short-term memory

0 引言

随着机动车数量快速增长，全球众多城市因交通拥堵、环境污染和能耗压力越来越大，开始积极地选择大力发展公共交通。公共自行车绿色、节能、环保，在满足"最初/最后一公里"、站点接驳、短途通勤等方面发挥着重要作用。公共自行车系统(public bike system，PBS)是单向交通系统，用户经常在不同站点租/还自行车，很容易出现站点库存失衡、需求满足率大幅下降的问题，因此必须提前预测短期需求，及时再平衡各站点库存，才能避免出现前述情况。可见，公共自行车短期需求预测是非常重要的公共交通研究课题。

按照空间粒度，可将公共自行车短期需求分为城市级、集群级和站点级[1]。城市级和区域级短期需求确定性强，预测难度较小，但预测结果不能直接作为站点库存再平衡的依据。站点级短期需求随机性强，预测难度较大，但预测结果可以直接用于再平衡站点库存。站点级预测模型包括统计模型和机器学习(Machine Learning，ML)模型两类。常见的统计模型有贝叶斯网络[2]、自回归移动平均[3]、多元回归[4, 5]、泊松和负二项式回归[6]等；传统的机器学习模型有支持向量机[7]、梯度提升树[8]、随机森林[9, 10]和神经网络[11, 12]等。统计模型更容易解释，传统的 ML 模型预测能力更强。不过，两者都需要人工识别隐藏在原始数据中的特征。

基于深度神经网络(deep neural network，DNN)的 ML 模型——深度学习(deep learning，DL)模型，不仅能自动提取隐藏特征[13]，而且学习和记忆能力更强，近年成为交通预测领域的主流模型。Zhang 等人[14]最早提出公共自行车短期需求预测 DL 模型，试图用卷积神经网络(convolution neural network，CNN)捕获用户需求的空间和时间依赖。基于欧式空间计算卷积是 CNN 擅长图像识别的重要原因，但公共自行车需求的空间依赖并非欧式关系，而且 CNN 没有记忆过去的机制，因此并不擅长捕获事务的时间依赖。

从循环神经网络(recurrent neural network，RNN)发展而来的长短期记忆(long short-term memory，LSTM)和门控循环单元(gated recurrent unit，GRU)，是专为捕获事务的时间依赖而设计的 DNN，最近几年受到了交通预测领域的关注。比如，Wang 等人[15]分别提出支持单时间步输出的公共自行车短期需求预测 LSTM 模型和 GRU 模型；Liu 等人[16]进一步提出支持多特性输入和多时间步输出的 LSTM 模型；许淼等人[17]针对出行分布的非严格周期性，提出引入注意力机制的共享单车短期需求预测 AM-LSTM 模型；刘耿耿等人[18]提出输出层与向前隐藏层和向后隐藏层同时连接的共享单车短期需求预测双向 LSTM 模型。LSTM、GRU 等 DNN 虽然

收稿日期：2022-01-20；**修回日期：**2022-03-14　**基金项目：**国家自然科学基金资助项目(71971171)

作者简介：乔健(1965-)，男(通信作者)，江苏宝应人，副教授，硕导，博士，主要研究方向为数据挖掘、交通管理(qiaojian@nwpu.edu.cn)；陈少博(2001-)，男，河南洛阳人，硕士研究生，主要研究方向为数据挖掘、交通预测；何梦莹(1999-)，女(回族)，宁夏平罗人，硕士研究生，主要研究方向为数据挖掘、交通预测.

擅长捕获事务的时间依赖，但不擅长捕获事务的空间依赖。

最近兴起的图卷积神经网络(GCNN)[19]基于非欧式空间计算图卷积，与公共自行车需求的空间依赖特征完全一致，因而取代 CNN 成为设计公共自行车短期需求预测模型的主流 DNN 之一。比如，在 Xiao 等人[20]提出的 STGCN 模型中，分别用 GCNN 和门 CNN(gated CNN)捕获用户需求的空间依赖和时间依赖；在 Kim 等人[21]提出的模型中，试图通过时、天和周三种时间粒度的 GCNN 的结合，同时捕获用户需求的空间和时间依赖，并将天气和时段(工作日/周末)等环境因素与全连接层输出的加权和作为模型的预测结果；在 He 等人[22]提出的模型中，用注意力机制与 GCNN 结合的多个图注意力卷积神经网络(GACNN)捕获不同时间粒度的车流，用一个全连接神经网络引入天气和兴趣点的影响。GCNN 虽然很符合用户需求在空间依赖上的非欧式特点，但与 CNN 一样，不擅长捕获用户需求的时间依赖。为此，Lin 等人[1]提出 GCNN_{rec}-DDGF 模型，分别用 GCNN 和 LSTM 捕获用户需求的空间和时间依赖；Chai 等人[23]提出 MGCNN 模型，用多个 GCNN 分别捕获站点之间的距离、交互和相关关系，用 LSTM 捕获多图卷积计算结果中的时间依赖。

用户需求受到很多环境因素的影响，其中有些因素是不变的，比如学校、车站、超市等兴趣点，有些因素是变化的，比如气温、降雨、风力等天气状况。事实上，不变环境因素是长期存在的，其影响已内化于用户需求，因此，构建预测模型时不应再考虑；可变环境因素是临时存在的，其影响并未内化于用户需求，因此，构建预测模型时应单独考虑，否则可能影响预测精度。目前，大多数模型都没有考虑环境因素，少数考虑了环境因素的模型又存在以下问题：一是忽视了不变和可变环境因素影响用户需求的性质差异，不必要地引入了一些不变环境因素，不仅重复，而且增加了模型的结构复杂度和时间复杂度；二是静态看待可变环境因素，忽视了这种环境因素的时间依赖性。

本文提出一种只考虑可变环境因素的 GCNN-LSTM-E 模型，用 GCNN 捕获用户需求的非欧式空间依赖，用 LSTM 捕获用户需求和可变环境因素的时间依赖，解决了现有模型存在的上述问题，提高了模型的预测性能。

1 GCNN-LSTM-E 模型设计

该模型由需求特征捕获器、环境特征捕获器和特征融合器组成，其架构见图 1。需求特征捕获器负责学习、记忆用户需求的变化规律，以及时间和空间依赖特性，环境特征捕获器负责学习、记忆可变环境因素的变化规律和时间依赖特性，特征融合器负责将可变环境因素与用户需求相融合，以反映前者对后者的影响。

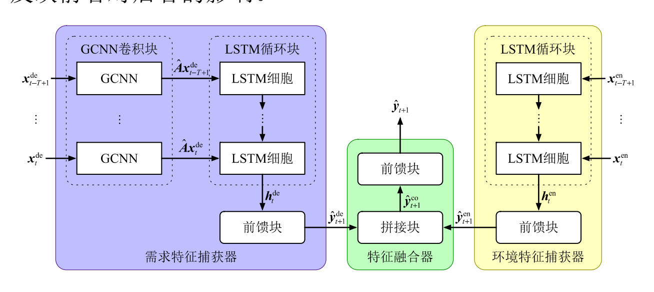


图 1 GCNN-LSTM-E 模型架构

Fig. 1 Architecture of the GCNN-LSTM-E model

1.1 需求特征捕获器

用户需求既有空间依赖性又有时间依赖性，且空间依赖是非欧式的，即站点间的 OD 关系是非对称的多对多关系，可以用图 $\boldsymbol{G}=(\boldsymbol{V},\boldsymbol{A})$ 表示这种关系，其中，$\boldsymbol{V}=\{v_1,\dots,v_N\}$ 表示站点集合，$\boldsymbol{A}=[a_{i,j}]\in\mathbb{R}^{N\times N}$ 表示邻接矩阵，$a_{i,j}$ 表示站点 v_i 与站点 v_j 之间租车和还车需求的相关度。假设用户需求为时间序列 $\boldsymbol{x}_{t-T+1}^{\text{de}},\dots,\boldsymbol{x}_t^{\text{de}}$，$\boldsymbol{x}_t^{\text{de}}=[x_{1,t}^{\text{de}},\dots,x_{N,t}^{\text{de}}]^T$ 表示 t 时刻 N 个站点的用户需求，$x_{i,t}^{\text{de}}$ 表示 t 时刻站点 v_i 的用户需求。

需求特征捕获器由 GCNN 卷积块、LSTM 循环块和全连接前馈块构成，用于捕获用户需求的空间和时间依赖。GCNN 卷积块负责通过图卷积运算 $\hat{\boldsymbol{A}}\boldsymbol{x}_{t-T+1}^{\text{de}},\dots,\hat{\boldsymbol{A}}\boldsymbol{x}_t^{\text{de}}$，从输入时间序列 $\boldsymbol{x}_{t-T+1}^{\text{de}},\dots,\boldsymbol{x}_t^{\text{de}}$ 中捕获用户需求的非欧式空间依赖，然后将运算结果输出到 LSTM 循环块。其中，$\hat{\boldsymbol{A}}$ 为图过滤器，通过下列步骤初始化[1]：1) 随机生成一个实对称邻接矩阵 $\boldsymbol{A}$；2) 令 $\boldsymbol{A}=\boldsymbol{A}+\boldsymbol{I}_N$，其中 $\boldsymbol{I}_N$ 是单位矩阵；3) 通过 $\hat{\boldsymbol{A}}=\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{D}^{-\frac{1}{2}}$ 计算出 $\hat{\boldsymbol{A}}$，其中 $\boldsymbol{D}\in\mathbb{R}^{N\times N}$ 是由 $d_{i,i}=\sum_j a_{i,j}$ 组成的对角矩阵。

LSTM 循环块负责从输入时间序列 $\hat{\boldsymbol{A}}\boldsymbol{x}_{t-T+1}^{\text{de}},\dots,\hat{\boldsymbol{A}}\boldsymbol{x}_t^{\text{de}}$ 中捕获用户需求的时间依赖。LSTM 细胞的结构如图 2 所示，其输出为短期状态 $\boldsymbol{h}_t^{\text{de}}$ 和长期状态 $\boldsymbol{c}_t^{\text{de}}$，由遗忘门 $\boldsymbol{f}_t^{\text{de}}$、输入门 $\boldsymbol{i}_t^{\text{de}}$ 和输出门 $\boldsymbol{o}_t^{\text{de}}$ 控制。遗忘门控制上一时刻的长期状态 $\boldsymbol{c}_{t-1}^{\text{de}}$ 中应该在当前时刻被遗弃的部分，输入门控制 $\tilde{\boldsymbol{c}}_t^{\text{de}}$ 中应该添加到 $\boldsymbol{c}_t^{\text{de}}$ 的部分，输出门控制 $\boldsymbol{c}_t^{\text{de}}$ 中应该输出到 $\boldsymbol{h}_t^{\text{de}}$ 的部分。

针对来自 GCNN 卷积块的输入 $\hat{\boldsymbol{A}}\boldsymbol{x}_{t-T+1}^{\text{de}},\dots,\hat{\boldsymbol{A}}\boldsymbol{x}_t^{\text{de}}$，LSTM 细胞用式(1)~(6)更新输入门 $\boldsymbol{i}_t^{\text{de}}$、遗忘门 $\boldsymbol{f}_t^{\text{de}}$、输出门 $\boldsymbol{o}_t^{\text{de}}$ 和细胞状态 $\boldsymbol{c}_t^{\text{de}}$ 的同时，计算短期状态序列 $\boldsymbol{h}_{t-T+1}^{\text{de}},\dots,\boldsymbol{h}_t^{\text{de}}\in\mathbb{R}^{d_1}$，从而捕获用户需求的时间依赖，这里的 d_1 表示隐藏单元的数量。

$$\boldsymbol{i}_t^{\text{de}}=\sigma(\boldsymbol{W}_i\cdot[\boldsymbol{h}_{t-1}^{\text{de}},\hat{\boldsymbol{A}}\boldsymbol{x}_t^{\text{de}}]+\boldsymbol{b}_i) \tag{1}$$

$$\boldsymbol{f}_t^{\text{de}}=\sigma(\boldsymbol{W}_f\cdot[\boldsymbol{h}_{t-1}^{\text{de}},\hat{\boldsymbol{A}}\boldsymbol{x}_t^{\text{de}}]+\boldsymbol{b}_f) \tag{2}$$

$$\boldsymbol{o}_t^{\text{de}}=\sigma(\boldsymbol{W}_o\cdot[\boldsymbol{h}_{t-1}^{\text{de}},\hat{\boldsymbol{A}}\boldsymbol{x}_t^{\text{de}}]+\boldsymbol{b}_o) \tag{3}$$

$$\tilde{\boldsymbol{c}}_t^{\text{de}}=\tanh(\boldsymbol{W}_c\cdot[\boldsymbol{h}_{t-1}^{\text{de}},\hat{\boldsymbol{A}}\boldsymbol{x}_t^{\text{de}}]+\boldsymbol{b}_c) \tag{4}$$

$$\boldsymbol{c}_t^{\text{de}}=\boldsymbol{f}_t^{\text{de}}\odot\boldsymbol{c}_{t-1}^{\text{de}}+\boldsymbol{i}_t^{\text{de}}\odot\tilde{\boldsymbol{c}}_t^{\text{de}} \tag{5}$$

$$\boldsymbol{h}_t^{\text{de}}=\boldsymbol{o}_t^{\text{de}}\odot\tanh(\boldsymbol{c}_t^{\text{de}}) \tag{6}$$

其中，$\boldsymbol{W}_i$、$\boldsymbol{W}_f$、$\boldsymbol{W}_o$、$\boldsymbol{W}_c$ 分别为输入门、忘记门、输出门和计算细胞状态时的权重矩阵，$\boldsymbol{b}_i$、$\boldsymbol{b}_f$、$\boldsymbol{b}_o$、$\boldsymbol{b}_c$ 分别为输入门、忘记门、输出门和计算细胞状态时的偏置项向量，$[\cdot,\cdot]$ 表示向量拼接，$\odot$ 表示逐元素相乘，$\sigma(\cdot)$ 和 $\tanh(\cdot)$ 分别表示 S 型函数和双曲正切函数。

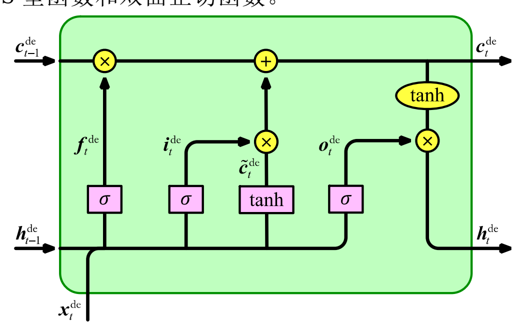


图 2 LSTM 细胞的结构

Fig. 2 Structure of a LSTM cell

全连接前馈块根据 LSTM 循环块的隐藏状态 $\boldsymbol{h}_t^{\text{de}}$ 计算输出 $\hat{\boldsymbol{y}}_{t+1}^{\text{de}}\in\mathbb{R}^N$，即用户需求的预测结果(引入可变环境因素前)。

1.2 环境特征捕获器

该部件由 LSTM 循环块和全连接前馈块构成，用于捕获可变环境因素的时间依赖。影响用户需求的环境因素主要包括兴趣点、天气状况和空气质量。兴趣点属于不变环境因素，天气状况和空气质量属于可变环境因素。本文认为，不变环境因素长期存在，其影响已内化于用户需求，气温、风力、降雨和空气质量等可变环境因素是临时的，其影响并未内化。为此，在模型中单独考虑这些因素。假设 $\boldsymbol{x}_t^{\text{en}}\in\mathbb{R}^N$ 表示 t 时刻的气温、风力、降雨和空气质量组成的向量。与需求特征捕获器用其 LSTM 循环块捕获用户需求的时间依赖类似，环境特征捕获器用其 LSTM 循环块接收输入时间序列 $\boldsymbol{x}_{t-T+1}^{\text{en}},\dots,\boldsymbol{x}_t^{\text{en}}$，并据此计算隐藏状态序列 $\boldsymbol{h}_{t-T+1}^{\text{en}},\dots,\boldsymbol{h}_t^{\text{en}}\in\mathbb{R}^{d_2}$，从而捕获可变环境因素的时间依赖，这里的 d_2 表示隐藏单元的数量。

全连接前馈块根据 LSTM 循环块的隐藏状态 $\boldsymbol{h}_t^{\mathrm{en}}$ 计算输出 $\hat{\boldsymbol{y}}_{t+1}^{\mathrm{en}} \in \mathbb{R}^M$，即可变环境因素的预测结果，其中 M 表示可变环境因素的数量。

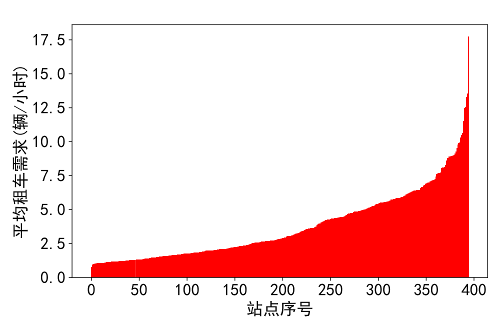
1.3　特征融合器

该部件由拼接块和全连接前馈块构成，用于将可变环境因素的影响施加在各站点的用户需求上。拼接块负责将各站点的用户需求的预测结果 $\hat{\boldsymbol{y}}_{i,t+1}^{\mathrm{de}}$ 与可变环境因素的预测结果 $\hat{\boldsymbol{y}}_{t+1}^{\mathrm{en}}$ 拼接成 N 个长度为 M+1 的向量 $\hat{\boldsymbol{y}}_{i,t+1}^{\mathrm{co}}=[\hat{\boldsymbol{y}}_{i,t+1}^{\mathrm{de}},\hat{\boldsymbol{y}}_{t+1}^{\mathrm{en}}]$，这些向量构成矩阵 $\hat{\boldsymbol{y}}_{t+1}^{\mathrm{co}}=[\hat{\boldsymbol{y}}_{1,t+1}^{\mathrm{co}},\ldots,\hat{\boldsymbol{y}}_{N,t+1}^{\mathrm{co}}]^T$，表示将 M 种可变环境因素的影响施加于 N 个站点的用户需求上。前馈块负责用公式 $\hat{\boldsymbol{y}}_{t+1}=\boldsymbol{W}_{\hat{y}^{\mathrm{co}}}\cdot\hat{\boldsymbol{y}}_{t+1}^{\mathrm{co}}$ 计算可变环境因素影响下的用户需求的预测结果 $\hat{\boldsymbol{y}}_{t+1}$，其中 $\boldsymbol{W}_{\hat{y}^{\mathrm{co}}}$ 是长度为 M+1 的权重向量。在模型训练阶段，用绝对值损失函数 $loss=\sum_{i=1}^{N}|\boldsymbol{y}_{i,t+1}-\hat{\boldsymbol{y}}_{i,t+1}|$ 计算 $\hat{\boldsymbol{y}}_{t+1}$ 与真实需求 $\boldsymbol{y}_{t+1}$ 间的误差，用 Adam 优化器[24]更新各权重矩阵和偏置项向量。为防止训练模型时过拟合，预先指定迭代阈值 s。在验证集上训练模型时，如果训练过程迭代 s 次后，预测精度仍未提高，就结束训练，并将此时的模型作为最终模型返回。

2　数值实验

2.1　数据准备

本文用纽约 PBS 的数据集验证本文模型。该数据集的时间跨度为 2017 年 1 月 1 日至 2019 年 12 月 31 日，包含 4 千多万条真实交易记录。本文对原始数据做如下预处理形成实际需求数据：先筛选出 3 年一直存在的站点，再统计这些站点每小时的平均租/还车需求，从中剔除每小时平均租/还车需求小于 1 个的站点后剩余 395 个站点，累计时长 26280 小时。图 3 所示为这些站点每小时平均租/还车需求分布情况，其中，左图为按递增顺序排列的租车需求分布，右图为与左图排列顺序一致的还车需求分布。可以看到，不同站点的需求变化很大，但同一站点的却非常接近。从后面的图 4 可以看到，3 种时间粒度的实际需求曲线均较平滑。图 3 和图 4 说明，绝大多数出行为往返交通且确定性很强。无同类竞争干扰是需求稳定的主要原因，也是本文选择该数据集的原因。



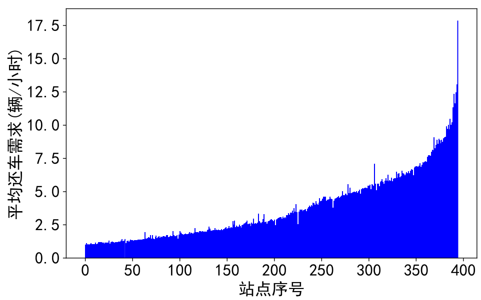


图 3　395 个站点的平均租/还车需求分布

Fig. 3　Distribution of average rental/return demands for 395 stations

表 1 为 395 个站点累计需求的描述性统计。根据租/还车需求均值可以推算出，平均每个站点每小时大约有 3.75 个租/还车需求。从中可以看出，租/还车需求的最小值和最大值差距很大，大多数站点的租/还车需求很小，一半以上的站点每小时平均租/还车需求不足 3.75 个。

表 1　累计需求的统计性描述

Tab. 1　Statistical description of cumulative demand

指标	租车需求	还车需求
均值	98787	98669
标准差	66880	67733
最小值	26429	20337
下四分位数	46324	45918
中位数	75682	74398
上四分位数	137882	138721
最大值	469548	465854

本文还用网络爬虫抓取了与需求数据时间段相同的纽约每天的气温、风力、降水量和空气质量数据，并将这些数据与不同时间粒度下的需求数据在时间上相对应。

2.2　评价指标

本文用均方根误差 *RMSE*(Root Mean Square Error)、平均绝对误差 *MAE*(Mean Absolute Error) 和判定系数 R^2(Coefficient of Determination)衡量模型的预测精度，它们的计算公式如下：

$$RMSE=\sqrt{\frac{1}{MN}\sum_{t=1}^{M}\sum_{i=1}^{N}(y_{i,t}-\hat{y}_{i,t})^2} \tag{7}$$

$$MAE=\frac{1}{MN}\sum_{t=1}^{M}\sum_{i=1}^{N}|y_{i,t}-\hat{y}_{i,t}| \tag{8}$$

$$R^2=1-\frac{\sum_{t=1}^{M}\sum_{i=1}^{N}(y_{i,t}-\hat{y}_{i,t})^2}{\sum_{t=1}^{M}\sum_{i=1}^{N}(y_{i,t}-\bar{y}_{i,t})^2} \tag{9}$$

其中，M 表示时长，N 表示站点数量，$y_{i,t}$、$\hat{y}_{i,t}$ 分别表示站点 i 在 t 时刻的预测值和真实值，$\bar{y}$ 表示平均值。*RMSE*、*MAE* 和 R^2 都反映了预测值与真实值之间的误差情况。对于相同的数据，用不同量纲表示时的 *RMSE*、*MAE* 和 R^2 是不同的。用同组数据预测时，*RMSE* 和 *MAE* 的值越小预测效果越好，R^2 的值越大预测效果越好。

2.3　超参数

本文用网格搜索法优化超参数，即人工指定超参数的搜索空间并穷举搜索。搜索空间用“开始:步长:结束”表示，比如“12:12:36”表示搜索空间为{12,24,36}。在验证数据集上，以 *MAE* 作为模型的评价指标，以确定超参数的最优值。GCNN-LSTM-E 模型有 3 类超参数：1) 需求特征捕获器和环境特征捕获器中循环块的时间步数 T、需求特征捕获器中 LSTM 细胞内的隐藏单元数 d_1、环境特征捕获器中 LSTM 细胞内的隐藏单元数 d_2；2) Adam 算法的学习率 α 和最小批处理规模 B；3) 模型训练过程的迭代阈值 s。

本文把需求数据集按时间顺序分为训练集、验证集和测试集，其中，验证集和测试集各 8 周，其余为训练集。然后用网格搜索法得到超参数的搜索空间、最优值及评价指标，如表 2 所示。

表 2　超参数的搜索空间、最优值及 GCNN-LSTM-E 模型的评价指标

Tab. 2　Search space and optimal value of the hyperparameters, as well as evaluation index of the GCNN-LSTM-E model

超参数						*MAE*
T	d_1	d_2	α	B	s	
12:12:36	25:25:200	25:25:200	.001:.001:.015	20:20:200	10:10:100	2.53
24	100	100	0.001	40	10	

2.4　实验结果

本文与 8 种基准模型进行了预测性能对比，这些模型分别为：HA(历史平均)模型、GCNN$_{\mathrm{rec}}$-DDGF(数据驱动的图卷积神经网络)模型、LASSO(带 l_1-范数惩罚的线性回归)模型、LSTM(长短期记忆)模型、MLP(多层感知器)模型、

SVR-linear(具有线性核函数的支持向量回归)模型、SVR-RBF(具有径向基核函数的支持向量回归)模型和 XGBoost(梯度提升回归树)模型。

9 种模型在 1h 时间粒度下租车需求测试集上实验的预测误差见表 3。可以看到，HA 模型因思想过于简单而性能表现最差；SVR-linear 模型和 LASSO 模型凭借回归技术显著提升了性能，但线性核函数导致其性能虽强于 MA 模型但弱于其他模型；SVR-RBF 模型采用非线性核函数，因而性能超越了 SVR-linear 模型和 LASSO 模型；MLP 模型凭借深度神经网络在性能上超越了 SVR-RBF 模型；LSTM 模型通过捕获需求序列的时间依赖使性能比 MLP 模型有所提升；XGBoost 模型通过支持微调和增加正则化参数以控制过拟合，而获得比 LSTM 模型更强的性能；$GCNN_{rec}$-DDGF 模型通过捕获用户需求的时空依赖，而获得比 XGBoost 模型明显更强的性能；GCNN-LSTM-E 模型不仅能捕获用户需求的时空依赖，而且考虑了可变环境因素的时间依赖，而获得比 $GCNN_{rec}$-DDGF 模型更强的性能。在还车需求测试集上的实验结果与上述类似，不再赘述。

表 3　在 1h 时间粒度下租车需求测试集上的预测误差

Tab. 3　Prediction errors on the rental demand test dataset under the time granularity of 1h

模型	*RMSE*	*MAE*	R^2
GCNN-LSTM-E	**2.80**	**1.71**	**0.73**
$GCNN_{rec}$-DDGF	2.89	1.73	0.72
XGBoost	3.31	1.98	0.65
LSTM	3.35	1.98	0.64
MLP	3.42	2.07	0.62
SVR-RBF	3.64	2.16	0.59
LASSO	3.68	2.27	0.58
SVR-linear	3.71	2.08	0.57
HA	4.69	2.86	0.34

GCNN-LSTM-E 模型和 $GCNN_{rec}$-DDGF 模型在 1h、0.5h 和 0.25h 时间粒度下租/还车需求数据集上的预测误差如表 4 所示。在 1h 时间粒度下，GCNN-LSTM-E 模型的 6 个指标均领先；在 0.5h 时间粒度下，GCNN-LSTM-E 模型的 4 个指标领先，2 个指标持平；在 0.25h 时间粒度下，GCNN-LSTM-E 模型的 2 个指标领先，1 个指标持平，3 个指标落后，说明 GCNN-LSTM-E 模型总体明显优于 $GCNN_{rec}$-DDGF 模型，引入可变环境因素的时间依赖，确提高了前者的预测性能。

从式(7)、式(8)和式(9)可知，*RMSE* 和 *MAE* 与自变量的量纲有关，R^2 与自变量的量纲无关。也就是说，不同时间粒度下的 R^2 能够直接比较，所以，可以确定表 4 中 1h 时间粒度下的 R^2 最好；但是，不同时间粒度下的 *RMSE* 和 *MAE* 不能直接比较，所以，无法确定表 4 中哪种时间粒度下的 *RMSE* 和 *MAE* 最好。

表 4　在 3 种时间粒度下租/还车需求测试集上的预测误差

Tab. 4　Experimental results on the rental/return demand test datasets under three kinds of time granularities

模型	时间粒度	租车需求			还车需求		
		RMSE	*MAE*	R^2	*RMSE*	*MAE*	R^2
GCNN-LSTM-E	1h	**2.80**	**1.71**	**0.73**	**2.80**	**1.69**	**0.75**
	0.5h	**1.93**	**1.28**	*0.66*	**1.91**	**1.28**	*0.69*
	0.25h	1.27	*0.85*	0.14	1.27	**0.83**	**0.21**
$GCNN_{rec}$-DDGF	1h	2.89	1.73	0.72	2.87	1.71	0.74
	0.5h	1.94	1.30	*0.66*	1.93	1.29	*0.69*
	0.25h	**1.26**	*0.85*	**0.15**	**1.26**	0.84	0.20

3 种时间粒度下纽约 PBS 的实际需求和 GCNN-LSTM-E 模型的预测需求如图 4 所示。可以看到，时间粒度越细，GCNN-LSTM-E 模型的预测精度越低，就是说，表 4 中 1h 时间粒度下的 *RMSE* 和 *MAE* 最好，预测精度最高。原因可能在于，时间单位越小，用户需求的随机性越强、波动性越大，导致模型的预测精度越低。因此，从兼顾平衡站点库存和提高预测精度的角度看，1h 是比 0.5h 和 0.25h 更合适的时间粒度。其他模型的实验结果与表 4 和图 4 的规律一致，不再赘述。

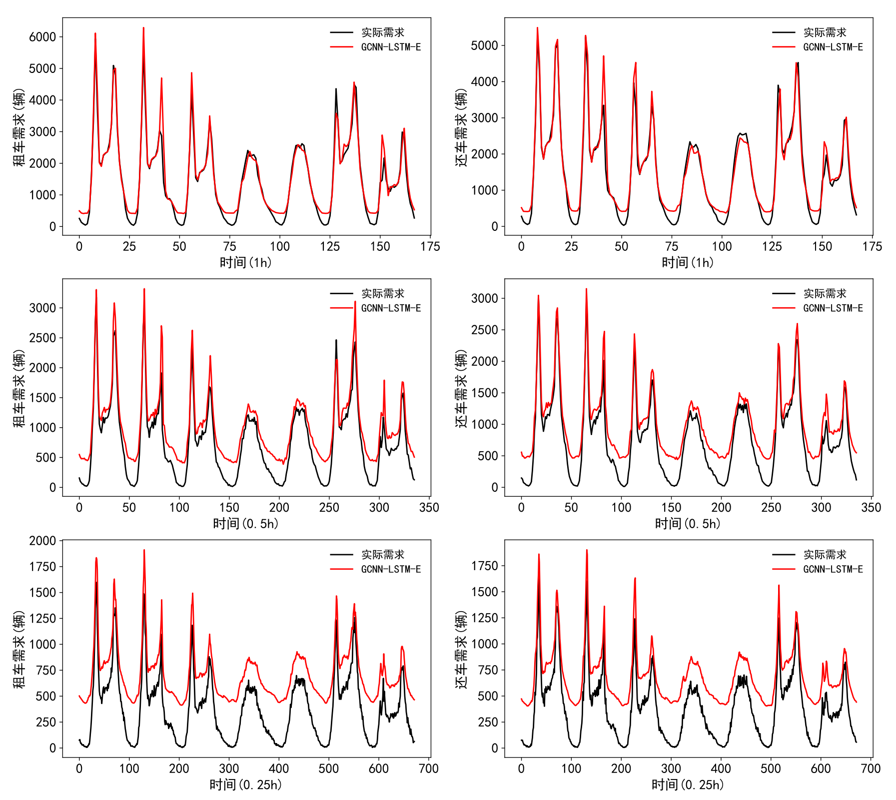


图 4　3 种时间粒度下纽约 PBS 的实际需求和 GCNN-LSTM-E 模型的预测需求

Fig. 4　Actual demand of New York PBS and predicted demand of the GCNN-LSTM-E model under three kinds of time granularities

3　结束语

公共自行车是一种单向交通工具，因此，站点库存再平衡对确保系统服务高水平运转非常关键。准确预测用户的短期需求是准确再平衡站点库存的前提条件，然而，大多数现有短期需求预测模型未考虑环境因素的影响，少数模型则忽视了不变和可变环境因素影响用户需求的性质差异、以及可变环境因素的时间依赖性。本文提出一种只考虑可变环境因素、用 GCNN 捕获用户需求的非欧式空间依赖、用 LSTM 捕获用户需求和可变环境因素的时间依赖、通过向量拼接和全连接网络将可变环境因素的影响施加于用户需求的 GCNN-LSTM-E 模型，并在 3 种时间粒度下，用真实的需求数据和可变环境数据进行了该模型与 8 种基准模型的预测性能对比。结果显示，在 1h 时间粒度下，GCNN-LSTM-E 模型的预测性能最优，而且优于 8 基准模型的预测性能，说明本文模型的设计合理有效。而且，从兼顾平衡站点库存和提高预测精度的角度看，1h 是最合适的时间粒度。

据了解，纽约 PBS 没有任何同类竞争者，而众多中国城市的 PBS 要面哈罗、美团、青桔等共享单车的激烈竞争。不同市场竞争环境下，用户需求的稳定性自然不同。本文模型仅在需求很稳定的纽约 PBS 数据集上进行了测试，今后有必要收集更多不同市场竞争环境下的 PBS 数据集，以对本文模型的普适性进行进一步的深入研究。此外，突发事件引发的需求突变也是模型应用时面临的难点。

参考文献：

[1] Lin L, He Z, Peeta S. Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach [J]. Transportation Research Part C: Emerging Technologies, 2018, 97: 258-276.

[2] Froehlich J E, Neumann J, Oliver N. Sensing and predicting the pulse of the city through shared bicycling [C]// Proc of the 21st International Joint Conference on Artificial Intelligence, Pasadena, 2009, 1420-1426.

[3] Kaltenbrunner A, Meza R, Grivolla J, *et al.* Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system [J]. Pervasive and Mobile Computing, 2010, 6 (4): 455-466.

[4] Rixey R A. Station-level forecasting of bikesharing ridership: Station network effects in three US systems [J]. Transportation Research Record, 2013, 2387 (1): 46-55.

[5] 朱才华，孙晓黎，李岩. 站点分类下的城市公共自行车交通需求预测 [J]. 吉林大学学报：工学版, 2021, 51 (02): 531-540. (Zhu Caihua, Sun Xiaoli, Li Yan. Forecast of urban public bicycle traffic demand by station classification [J]. Journal of Jilin University: Engineering and Technology Edition, 2021, 51 (02): 531-540.)

[6] Rudloff C, Lackner B. Modeling demand for bikesharing systems: neighboring stations as source for demand and reason for structural breaks [J]. Transportation Research Record, 2014, 2430 (1): 1-11.

[7] Xu H T, Ying J, Wu H, *et al.* Public bicycle traffic flow prediction based on a hybrid model [J]. Applied Mathematics & Information Sciences, 2013, 7 (2): 667-674.

[8] Chen T, Guestrin C. Xgboost: A scalable tree boosting system [C]// Proc of the 22nd acm sigkdd international conference on knowledge discovery and data mining, San Francisco: ACM, 2016, 785-794.

[9] Yang Z, Hu J, Shu Y, *et al.* Mobility modeling and prediction in bike-sharing systems [C]// Proc of the 14th Annual International Conference on Mobile Systems, Applications, and Services, New York: ACM, 2016, 165-178.

[10] 陈红，陈恒瑞，史转转，等. 公共自行车使用时空特性挖掘及租还需求预测 [J]. 交通运输系统工程与信息, 2021, 21 (02): 238-244+250. (Chen Hong, Chen Hengrui, Shi Zhuanzhuan, *et al.* Spatiotemporal characteristics mining and demand forecasting of shared bicycle borrow and return [J]. Journal of Transportation Systems Engineering and Information Technology, 2021, 21 (02): 238-244+250.)

[11] Chang P C, Wu J L, Xu Y, *et al.* Bike sharing demand prediction using artificial immune system and artificial neural network [J]. Soft Computing, 2017: 1-14.

[12] 解小平，邱建东，汤旻安. 基于 Elman 神经网络的公共自行车单站点需求预测 [J]. 计算机工程与应用, 2017, 53 (16): 221-224+236. (Xie Xiaoping, Qiu Jiandong, Tang Minan. Demand prediction of public bicycle rental station based on Elman neural network [J]. Computer Engineering and Applications, 2017, 53 (16): 221-224+236.)

[13] Lecun Y, Bengio Y, Hinton G. Deep learning [J]. Nature, 2015, 521: 436.

[14] Zhang J, Zheng Y, Qi D, *et al.* DNN-based prediction model for spatio-temporal data [C]// Proc of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, New York: ACM, 2016, 1-4.

[15] Wang B, Kim I. Short-term prediction for bike-sharing service using machine learning [C]// Proc of the International Symposium of Transport Simulation, Amsterdam: Elsevier, 2018, 171-178.

[16] Liu X, Gherbi A, Li W, *et al.* Multi features and multi-time steps LSTM based methodology for bike sharing availability prediction [C]// Proc of the 16th International Conference on Mobile Systems and Pervasive Computing, Amsterdam: Elsevier, 2019, 394-401.

[17] 许淼，刘宏飞，初凯. 基于 AM-LSTM 模型的共享单车时空需求预测 [J]. 湖南大学学报：自然科学版, 2020, 47 (12): 77-85. (Xu Miao, Liu Hongfei, Chu Kai. Spatiotemporal demand prediction of bike-sharing based on AM-LSTM model [J]. Journal of Hunan University: Natural Sciences, 2020, 47 (12): 77-85.)

[18] 刘耿耿，朱予涵，郭灿阳. 基于双向长短期记忆网络的共享单车流量预测 [J]. 小型微型计算机系统, 2021, 42 (09): 1871-1876. (Liu Genggeng, Zhu Yuhan, Guo Canyang. Sharing bicycle flow forecasting based on bi-directional long short-term memory [J]. Journal of Chinese Computer Systems, 2021, 42 (09): 1871-1876.)

[19] 徐冰冰，岑科廷，黄俊杰，等. 图卷积神经网络综述 [J]. 计算机学报, 2020, 43 (05): 755-780. (Xu Bingbing, Cen Keting, Huang Junjie, *et al.* A survey on graph convolutional neural network [J]. Chinese Journal of Computers, 2020, 43 (05): 755-780.)

[20] Xiao G, Wang R, Zhang C, *et al.* Demand prediction for a public bike sharing program based on spatio-temporal graph convolutional networks [J]. Multimedia Tools and Applications, 2020, 80 (15): 22907-22925.

[21] Kim T S, Lee W K, Sohn S Y. Graph convolutional network approach applied to predict hourly bike-sharing demands considering spatial, temporal, and global effects [J]. Plos One, 2019, 14 (9): 1-16.

[22] He S, Shin K G. Towards fine-grained flow forecasting: A graph attention approach for bike sharing systems [C]// Proc of The Web Conference 2020, New York: ACM, 2020, 88-98.

[23] Chai D, Wang L, Yang Q. Bike flow prediction with multi-graph convolutional networks [C]// Proc of the 26th ACM SIGSPATIAL international conference on advances in geographic information systems, New York: ACM, 2018, 397-400.

[24] Kingma D P, Ba J. Adam: A method for stochastic optimization [Preprint]. arXiv preprint arXiv: 1412. 6980, 2014.